

Classification improvement of spoken arabic language based on radial basis function

Thamir Rashed Saeed, Jabar Salman, Alaa Hussein Ali

Digital System Design and Pattern Recognition Research Group (DSDPRG), Smart Sensor System, Radar Signal Processing, Pattern Recognition, Department of Electrical Engineering, University of Technology, Iraq

Article Info

Article history:

Received Feb 8, 2018

Revised Jul 26, 2018

Accepted Aug 15, 2018

Keywords:

Arabic language letters

Classification

Radial basis function

ABSTRACT

The important task in the computer interaction is the languages recognition and classification. In the Arab world, there is a persistent need for the Arabic spoken language recognition To help those who have lost the upper parties in doing what they want through speech computer interaction. While, the Arabic automatic speech recognition (AASR) did not receive the desired attention from the researchers. In this paper, the Radial Basis Function (RBF) is used for the improvement of the Arabic spoken language letter. The recognition and classification process are based on three steps; these are; preprocessing, feature extraction and classification (Recognition). The Arabic Language Letters (ALL) recognition is done by using the combination between the statistical features and the Temporal Radial Basis Function for different letter situation and noisy condition. The recognition percent are from 90% - 99.375% has been gained with independent speaker, where these results are over-perform the earlier works by nearly 2.045%. The simulation has been made by using Matlab 2015b.

Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Thamir Rashed Saeed,
Department of Electrical Engineering,
University of Technology, Iraq.
Email: thamir_rashed@yahoo.com

1. INTRODUCTION

The recognition and classification of the Arabic Langue Letter is the interesting subject in the applications of Arabic computer interaction. The computer interaction is an important tool in the intelligent systems and technologies. The Language recognition is speech recognition, and it is characterized as the way toward changing over sound waves (acoustic discourse signals) to its relating set of words or other linguistic units [1]. In this context, that recognition is based on a specific algorithm step, where these algorithms are based on the feature extraction of the selected subject which is required to recognize it, while the features represent the carrier of the speaker essence [2], [3]. Where, these features will be reduced to minimize the efforts of digital signal processing applications [4], [5], [6], [7].

In this context, the speech signal also carries the information of the particular speaker, including social factors, affective factor and the properties of the real voice production [8]. In effect, the speech has the potential of being an important mode of interaction with the computer. Speech processing is one of the exciting areas of signal processing [9]. The letters of the Arabic language are different from the rest of the languages because the letter pronunciation is differed according to their position in the word. Also, their pronunciation varies according to the impact of the word in the sentence. As well the letter pronunciation is between the Arab countries according to their dialect.

The recognition process has been based on two phases, training, and testing. Where, the training phase work with extracted features by using suitable neural network (NN) algorithm, and then use this algorithm in testing phase. There are many NN algorithms types, the Radial Basis Function (RBF) is one of

the optimal algorithms in a noisy environment. In this context, this algorithm is a linear combination of radial basis functions and can use in a function approximation, time series prediction, and control. The advancement of this algorithm over the other is the faster convergence, smaller extrapolation errors and higher reliability [10].

Since the last years, the researchers have been looking into optimal features and ways to recognize the Arabic Language letters [11]. In this context and for the importance of the subject there are many types of research in this area, some of these;

In [12] a Hidden Markov Model (HMM) is used as feature extraction algorithm, while the noise reduction is made by using power spectral estimator, and the Gabor filter bank is used for the noise separation in an acoustic event detection system. However, in [13] was used HMM with Mel frequency Cepstral Coefficient (MFCC) features under no noise condition for speech recognition. While in [14] there are five speech parameters have been used as features for speech recognition. These parameters are; Relative Spectra Processing (RASTA), MFCC, Linear Predictive Coding (LPC) Analysis, Dynamic Time Wrapping (DTW) and Zero Crossings with Peak Amplitudes (ZCPA). Where, RASTA and MFCC are Extracted as features In addition to being factors, While LPC predicts as features based on previous features. In [15], a Cepstral frequency coefficient and perceptual linear prediction have been presented as feature extraction methods. While, Rasta filtering and Cepstral mean subtraction has presented as feature normalization technique, with a combination of Gaussian mixture models (GMM) and linear/non-linear kernels which is based on support vector machine (SVM) as speaker identification. In [16] an FFT is used as gender identification, then use back-end system to create a gender model to recognize these genders with an average of the accuracy 80%. The use of signal processing technique for speech recognition for a particular language is presented in [17], while the feature extraction is based on the adopted algorithms. Also, the comparison between these adopted algorithms is presented in [18]. A hybrid of HMM and Radial Basis Function (RBF) was presented for continuous speech recognition with 65% recognition rate in [10].

The problem lies, in addition to the lack of interest in the research in the Arabic automatic speech recognition, the most of the published papers dealing with an HMM algorithm. Where the accuracy of ASR using HMM algorithm is affected by several factors; the phoneme set used; the number of HMM states allocated for each phoneme and the duration of each phoneme, in addition to the noisy environments, thus reducing this accuracy.

Therefore, this paper gives an overview of speech features extraction and the proposed work which is consisting of three steps; preprocessing, feature extraction, classification and finally the comparison with other works. This study has been dealt with differently letter position in the word, letter's impact of the word in the sentence and letter pronunciation of different Arab country's dialect and. Also, the number of training pattern has increased from 10-30 per class with a constant testing pattern for each class with the noisy condition for recognition the independent speakers. Also, the proposed work has been based on the Radial Basis Function Neural network with statistical features. Then the comparison has been made among the previous works which are used HMM and other algorithms.

2. SPEECH FEATURE EXTRACTION

The structure of the vocal organs generates a wide variety of waveforms. These waveforms can be broadly categorized into voiced and unvoiced speech; this categorization is made after the features extraction [9]. In this context, two kinds of algorithms which are used for feature extraction; the first one is related to speech processes, while the second is related to the results of these processes. While, the feature vectors are equivalent to the vectors of explanatory variables used in statistical procedures such as linear regression [19]. Therefore, the features are;

a. Articulatory features

Articulatory features (AFs) have attracted interest from the speech recognition community, where, these features describe the configuration of the human vocal tract and the properties of speech products. The essential thought of this approach is to bear a proclivity to the articulatory occasions fundamental the discourse flag. This portrayal is made out of classes depicting a basic articulatory properties of discourse sounds, for example, put, way, voicing, lip adjusting, the opening between the lips, and the position of the tongue.

b. Features based on perception system

The auditory system has been based on the sensory system for the feeling of hearing. The research in speech recognition is dealing with the way in which the human can recognize the speech and use the speech information to understand the spoken language [20]. In effect, the statistical features can be considered as the second kind of features, but it's related to the first kind also. Therefore the statistical feature can represent as an active feature in the speech recognition application.

3. PROPOSED WORK

The proposed work in this paper consists of three stages; preprocessing, statistical feature extraction, and classification.

a. Preprocessing

The preprocessing stage is represented as preparing stage, where it prepares the signal to the feature extraction stage. Therefore, this stage consists of five steps; these are; Saliency is removed, Normalized, pre-emphasis, Framing and windowing, and then take one frame. In this context, the saliency removing is done to reduce the size of data which need to process and keep the samples which contain the information only. The normalization step is a limitation of the sample values which need to process it. Pre-emphasis is a signal concentrated step and boosted the energy at the upper band frequency, while the framing is segmented step also to reduce the process time and data size. One Arabic letter has been taken as an example in this stage as in Figure 1. Where, the sound signal of a sad (ص) letter from Arabic alphabet letters has been taken in real environments, as in Figure (1a-b), then removing the effect of the environment. The preprocessing steps results have further strengthened our confidence in the statistical features as the classification tools, where the saliency is removed, and normalization then framing and select one frame with a window which gives signal effect with decreasing the number of process samples as shown in the Figure 1.

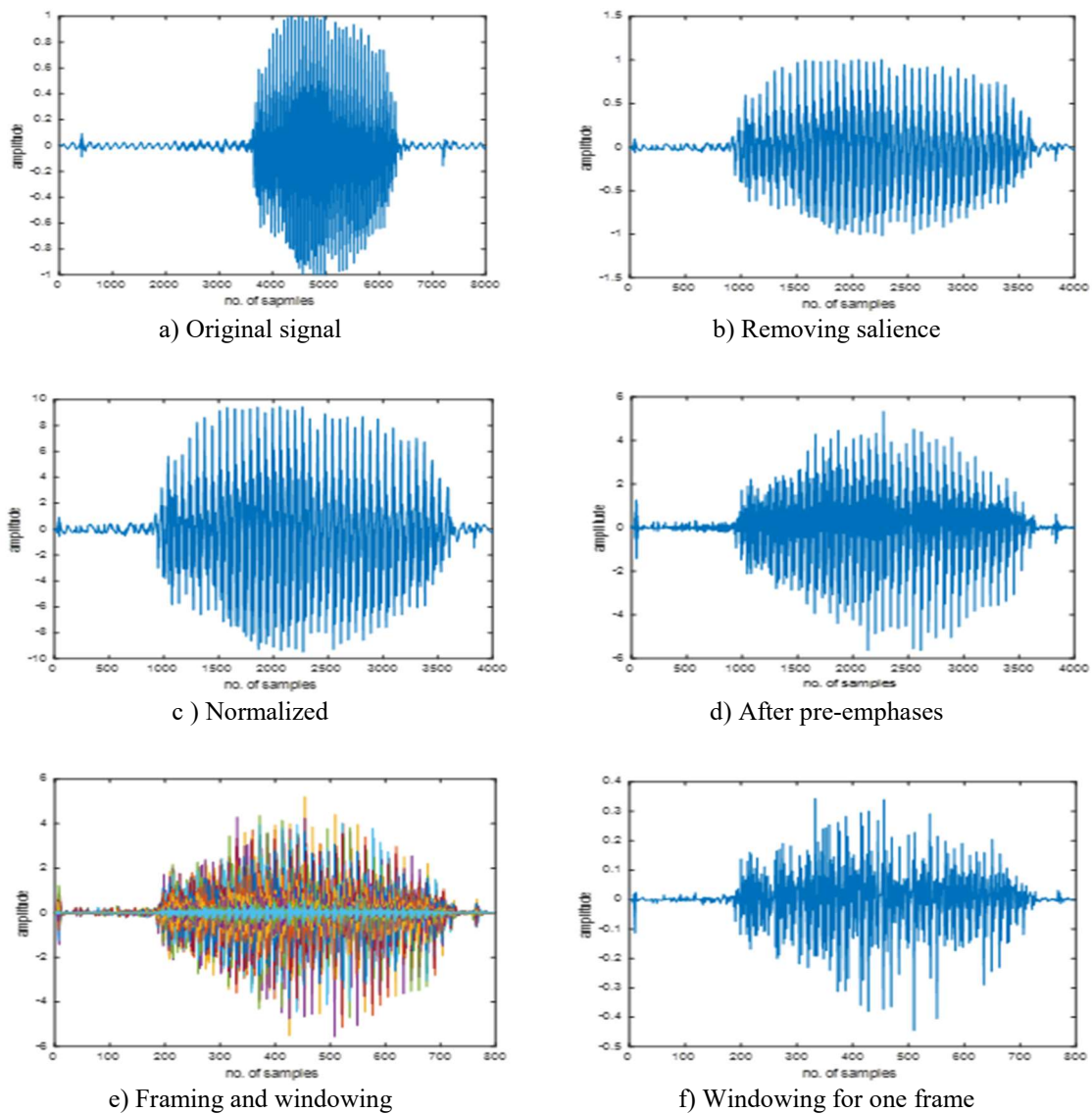


Figure 1. Preprocessing stage

b. Statistical Features

Statistical features have represented the core of the signals, and It carries the spirit of the signal. Where some of these features are; zero crossing rate, signal Energy, temporal centroid, d) energy entropy (EE), RMS, spectral flux, Spectral energy, and MFCC. Where these features have been representing the suitable features for the sound signal as in Figure 2. After preprocessing step, some clear-cut, effectively feature was used instead of all extracted feature to reduce processing time while maintaining the accuracy.

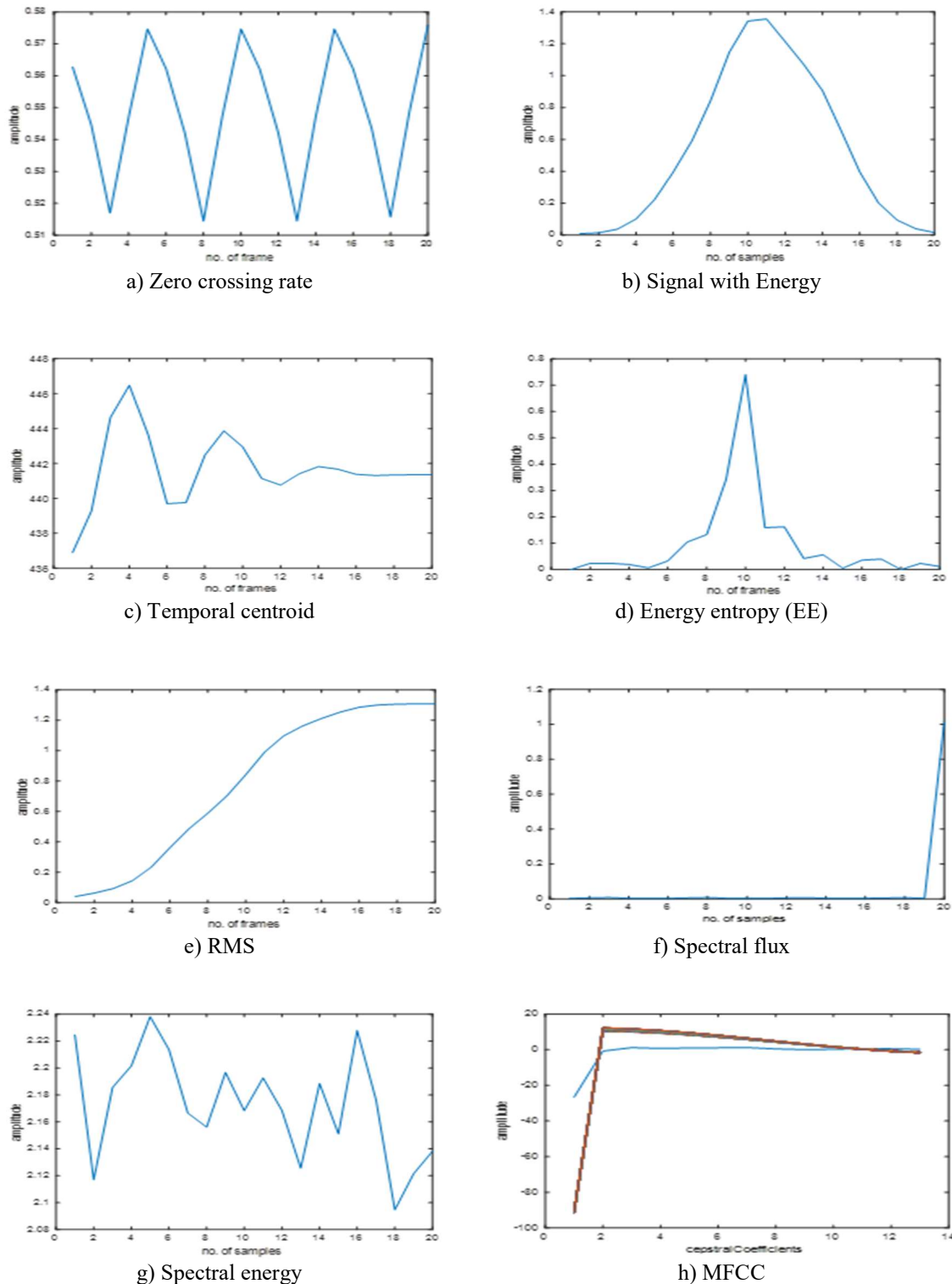


Figure 2. Statistical features of the sound signal (for Sad [ص] letter)

c. Classification

The classification stage has been based on the Radial Basis Function(RBF) neural network as shown in Figure 3. This stage is affected by many factors, one of the most powerful ones is the number of training and testing patterns. Therefore, the increasing of the training pattern will cause to undermine the similarity between the patterns to appear the difference between these patterns. Many experiments have been done with different letter position in the word, letter's impact of the word in the sentence and letter pronunciation of different Arab country's dialect. In this context, the number of training pattern has also increased from 10-30 per class with ten testing pattern for each class. Tables 1, 2, 3, and 4. shown the results of classification of different experimental parameters.

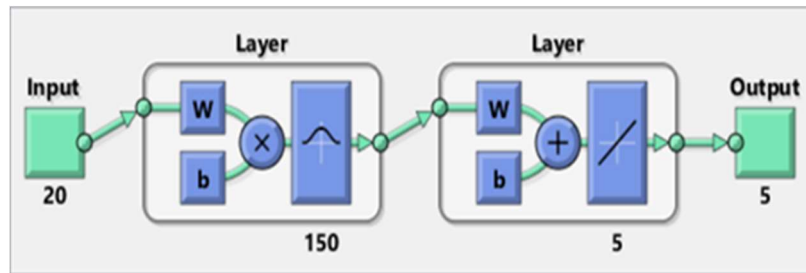


Figure 3. Radial Basis Function(RBF) neural network

Table 1. Recognition of Bee (ب) Letter with a Different Number of Training Patterns for Different Letter position, impact and letter pronunciation of a different Arab country's dialect

Letters	No of training	Results %	No of training	Results %	No of training	Results %
ب	10	50	20	80	30	90
ج	10	0	20	20	30	0
س	10	10	20	0	30	0
ك	10	0	20	0	30	0
ص	10	40	20	0	30	10

Table 2. Recognition of Bee (ب) Letter with a Different Number of Training Patterns for one Arab Country's Dialect with Different Letter Position and Impact

Letters	No of training	Results %	No of training	Results %	No of training	Results %
ب	10	91.66	20	96.66	30	98.33
ج	10	0	20	3.33	30	0
س	10	1.66	20	0	30	0
ك	10	0	20	0	30	0
ص	10	6.66	20	0	30	1.66

Table 3. Recognition of Bee (ب) Letter with a Different Number of Training Patterns for One Arab Country's Dialect with Different Letter Position

Letters	No of training	Results %	No of training	Results %	No of training	Results %
ب	10	96.875	20	98.875	30	99.375
ج	10	0	20	1.25	30	0
س	10	0.625	20	0	30	0
ك	10	0	20	0	30	0
ص	10	2.5	20	0	30	0.625

Table 4. Classification by Using the Temporal Radial Basis Function(TRBF) Neural Network

Letters	ص	ك	س	ج	ب
ب	99.375%	0	0	0.625	0
ج	1.25	98.125%	0.625	0	0
س	0	0.625	98.125%	1.25	0
ك	1.875	0	00.625	97.5%	0
ص	1.25	0	0	0	98.75%

The average classification is 95.9% for five letters in experimental which was conducted with different parameters of letters situations, while the classification is 98.175% for one Arab country's dialect with different letter position. Therefore, the advancement of the proposed algorithm has been proved by the comparison between the results (classification) which has been gained with other work as presented in the Table 5.

Table 5. Comparison proposed work with others

Ref	Classification method	Average Recognition %
[12]	HMM	74.5%
[21]	TMNN	90.7%
[22]	MLP	96.3%
Previous work[23]	MLFFNN	96.33%
Present work	TRBFNN	98.175%

4. CONCLUSION

The recognition which is based on the combination of the statistical features with the Radial Basis Function (RBF) as the recognition neural network algorithm are gaining an overperform the other combinations by 1.845%. This advancement of that combination is caused by using RBF where in this algorithm the hidden function is a Gaussian, while the Euclidean distance is computed from the test point to the main center of each neuron. Therefore, the average recognition rate, which is a gain of that combination is 98.175%. Also, the parameters which are effected on the Arabic letter classification are letter position in the word, letter's impact of the word in the sentence and letter pronunciation of different Arab country's dialect.

ACKNOWLEDGEMENTS

We would like to thank Asst. prof. Dr. Mahmmod M Hamza for his support and advice in accomplishing our work.

REFERENCES

- [1]. Khalid M.O, Nahar, Moustafa, Elshafei, Wasfi G, "Al-Khatib and Husni Al-Muhtaseb. Statistical Analysis of Arabic Phonemes for Continuous Arabic Speech Recognition," *International Journal of Computer and Information Technology*, (ISSN: 2279 – 0764), vol. 01 no. 02, November 2012.
- [2]. Srinivasa Perumal R, Nadesh R. K. and Senthil Kumar N. C., "Robust Face Recognition Using Enhanced Local Binary Pattern," *Bulletin of Electrical Engineering and Informatics*, vol. 7, no. 1, pp. 96-101, March 2018.
- [3]. Shofwatul 'Uyun and Lina Choridah. "Feature Selection Mammogram based on Breast Cancer Mining," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 1, pp. 60-69, February 2018.
- [4]. Aws Al-Qaisi, Saleh A. Khawatreh, Ahmad A. Sharadqah, Ziad A. Alqadi, "Wave File Features Extraction Using Reduced LBP", *International Journal of Electrical and Computer Engineering (IJECE)*, vol 8, no. 5, Oct., part1, 2018.
- [5]. Hussein Ali Aldelfy, Mahmood Hamza Al-Mufraji, Thamir R. Saeed, "Improved Key Frame Extraction Using Discrete Wavelet Transform with Modified Threshold Factor," *Telecommunication Computing Electronics and Control (TELKOMNIKA)*, vol.16, no.2, pp. 567-572, April 2018.
- [6]. Hameed R. Farhan, Mahmud H. Al-Muifraje and Thamir R. Saeed. A Novel Face Recognition Method based on One State of Discrete Hidden Markov Model," *IEEE Annual Conference on New Trends in Information & Communications Technology Applications*, pp. 7-9, March 2017.
- [7]. Hameed R. Farhan, Mahmud H. Al-Muifraje and Thamir R. Saeed, "Using only Two States of Discrete HMM for High-Speed Face Recognition," *IEEE Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications*, IRAQ, pp. 9-10, May, 2016.
- [8]. Sadek Ali. "Gender Recognition System Using Speech Signal," *International Journal of Computer Science, Engineering and Information Technology (IJCEIT)*, vol. 2, no.1, February 2012.
- [9]. Satyanand Singh. "Forensic and Automatic Speaker Recognition System," *IJECE*, vol. 8, no. 5, Oct., part1, 2018.
- [10]. Judith Justin and 2Ila Vennila. "A Hybrid Speech Recognition System with Hidden Markov Model and Redial Basis Function Neural Network," *American Journal of Applied Sciences*, vol. 10, no.10, 2013.
- [11]. Megha Agrawal and Tina Raikwar, "Speech Recognition Using Signal Processing Techniques," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 5, no. 8, February 2016.
- [12]. Jens Schro"der, Benjamin Cauchi, Marc Rene' Scha"dlar, Niko Moritz, Kamil Adiloglu, Jo"rn Anemu"ller, Simon Doclo1, Birger Kollmeier and Stefan Goetze, "Acoustic Event Detection Using Signal Enhancement and Spectro-Temporal Feature Extraction," *IEEE, Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, 2013.

- [13]. Shaik Shafee and Prof.B.Anuradha, "Isolated Telugu Speech Recognition using MFCC and Gamma Tone Features by Radial Basis Networks in Noisy Environment," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, no. 3, March 2015.
- [14]. Pratik K. Kurzekar, Ratnadeep R. Deshmukh, Vishal B. Waghmare, Pukhraj P. Shrishrimal, "A Comparative Study of Feature Extraction Techniques for Speech Recognition System," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 12, December 2014.
- [15]. Imen TRABELSI, Dorra BEN AYED, "On the Use of Different Feature Extraction Methods for Linear and Non Linear kernels," *IEEE Sciences of Electronics, Technologies and Telecommunications*, March 21-24, Tunisia, 2012.
- [16]. Sadek Ali, Shariful Islam and Alamgir Hossain, "Gender Recognition System Using Speech Signal," *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*, vol.2, no.1, February 2012.
- [17]. Megha Agrawal, Tina Raikwar, "Speech Recognition Using Signal Processing Techniques," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 5, no. 8, February 2016.
- [18]. M. Honda, "Human Speech Production Mechanisms," *NTT Technical Review*, vol. 1, no. 2 May 2003.
- [19]. Liu, H., Motoda H, "Feature Selection for Knowledge Discovery and Data Mining," Kluwer Academic Publishers, Norwell, MA, USA. *The Springer International Series in Engineering and Computer Science book series (SECS)*, vol. 454, 1998.
- [20]. D. Kim, S. Lee, "Auditory Processing of Speech Signals for Robust Speech Recognition on Real World Noisy Environment," *IEEE Transaction on Speech and Audio Processing*, vol. 7, no. 1, 1999.
- [21]. Khalid Saeed, "A New Feature Extraction Method for TMNN-Based Arabic Character Classification," *Computing and Informatics*, vol. 26, pp. 403-4020, 2007.
- [22]. Manal El-Obaid, Amer Al-Nassir and Iman Abuel Maaly, "Arabic Phoneme Recognition using Neural Networks", *Proceeding SIP'06 Proceedings of the 5th WSEAS international conference on Signal processing*, pp. 99-104, Istanbul, Turkey-May, pp. 27-29, 2006.
- [23]. Jabar Salman, Alaa Hussein Ali and Thamer Rashed Saeed, "Improve the Recognition of Arabic Sign Languages Based on Statistical Features," *Accepted in Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE)*, 2018.

BIOGRAPHIES OF AUTHORS



Thamer Rashed Saeed was Born in Baghdad, Iraq, on February 10, 1965. He received the B.Sc. Degree from military engineering college in Baghdad in 1987, the M.Sc. Degree from military engineering college in Baghdad in 1994 and Ph.D. degree from AL-Rashed college of engineering and Secinec in Baghdad 2003. From 1994 to 2003, he worked with military engineering college in Baghdad as a member of teaching staff. From 2003 till now, he worked with the University of Technology in Baghdad as a member of teaching staff. Currently, he is the Asst. Professor of electrical engineering at university of Technology and a head of radar research group in the Electrical Eng. Dept.. His major interests are in digital signal processing, digital circuit design for DSP based on FPGA, sensor network and Pattern Recognition.



Jabbar Slman Hussain has born in 1974 in Baghdad, Iraq and he received the B.Sc. Degree in electronic and communication engineering in 1998 from university of technology, department of electrical and electronic engineering, Baghdad, Iraq. He obtained his M.Sc Degree in electronics engineering from university of technology, department of electrical and electronic engineering, Baghdad, Iraq, in 2001. He is working in university of technology, Baghdad, Iraq, as a lecturer in the Department of Electrical Engineering, electronic branch. In the research field, is the active member of Digital System Design and Pattern Recognition Research Group (DSDPRG)



Dr. Alaa Hussein Ali Al-Ameri was Born in Baghdad, Iraq on 1970. He received the B.Sc. Degree from military engineering college in Baghdad in 1993, the M.Sc. Degree from military engineering college in Baghdad in 2002 and Ph.D. degree from AL-Rashed college of engineering and Science in Baghdad 2008. From 2003 till now, he worked with the University of Technology in Baghdad as a member of teaching staff. Currently, he is the Asst. Professor of electrical engineering at university of Technology.